DOCUMENT RESUME

ED 158 639 HE 010 312

AUTHOR Cox, Roy
TITLE Formative Evaluation: Interpretation and Participation.
INSTITUTION London Univ. (England). Inst. of Education.
PUB DATE [74]
NOTE 14p.; This paper has been reproduced from "Contemporary Issues in Educational Testing" edited by Hans Crombag and Daton de Gruyter, Mouton, The Hague
AVAILABLE FROM University of London Institute of Education, UTMU, 55 Gordon Square, London WC1H ONT, England

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Course Evaluation; Educational Improvement; Evaluation Methods; *Evaluation Needs; *Formative Evaluation; Higher Education; Informal Assessment; Student Attitudes; *Student Behavior; *Student Evaluation; Student Motivation; Student Participation; *Student Responsibility; Teaching Methods
IDENTIFIERS Great Britain

ABSTRACT

An Approach to formative evaluation is described that attempts to actively involve students and explore the meaning of processes, activities, and performances in terms of the wide variety of students' perceptions and responses. When poor student performance is not attributable to lack of effort or ability, the need for a more personal understanding of the meaning of differential formative evaluation is important. Student evaluation needs to be seen within the broader social context of education and especially in relation to the psychology of adolescence. Two individual cases are discussed to illustrate this viewpoint. Different responses of students to particular types of assessment and problems concerning the interpretation of grades are covered. Research from Edinburgh on "cue-consciousness" discusses student expectations and performance and the importance of such behavior in understanding the meaning of formative evaluation. An evaluation study of a course for teachers in Medical General Practice is described to illustrate the need of students for both independence and freedom. The role of the student in course evaluation was not as consumer but as participant who had the responsibility for making the course valuable. Problems that arise in evaluation of student projects or guided work for some content areas make it important to explore individual responses and involve the students more. Student involvement in assessment anchors the learner more in the characteristics of the work rather than in the authority of the teacher. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED158639

UTMU
University Teaching Methods Unit

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Univ of London
Dept of Education

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

University Teaching Methods Unit
University of London Institute of Education
55 Gordon Square London WC1H 0NT

HE 010 312

FORMATIVE EVALUATION:
INTERPRETATION AND PARTICIPATION

ROY COX

30P

This paper has been reproduced from "Contemporary Issues in Educational Testing" edited by Hans Crombag and Daton de Gruyter, Mouton, The Hague, by kind permission of the publishers

Introduction

Measures of achievement in formation evaluation provide us with both 'answers' and with 'problems'. The 'answers' are generally thought of as useful in both improving the teaching and the learning; the problems are often ignored or referred to the counsellor or the psychiatrist. We tend to see the average scores on our differential tests as answering our questions about where the teaching needs to be strengthened or modified. The individual scores tell us where particular students should divert their efforts - what sort of exercises and extra reading might be necessary. This is a very important area and one in which Britain is generally backward compared with the United States. 'Remedial' is a word which is only just emerging from the primary school and is rare and suspect in higher education. The problems arise when we feel lack of achievement is not simply lack of effort - to be remedied by more general or selective effort, *or* lack of ability - to be remedied by more effort or perhaps a less demanding course. But then there are people employed to look after those awkward cases and it is their job to bring these students back into the fold. (The fact that some may no longer see their job quite like that is a rather distant problem.) But is this a tenable division of labour? Is there an area of interest between these rather formal definitions of the roles of teacher and counsellor which is under-emphasised?

Measures of student opinion and satisfaction in evaluating courses provide us with some answers but usually many more problems. Both the answers and the problems - if we can solve them - can help us to improve our courses and our teaching in both a formative and a summative sense. Again the United States is generally far ahead of Britain in the design and use of course evaluation techniques. But is it true that in this case the problems are primarily our problems? The students provide *us* with information to work on to provide *them* with a better course? Perhaps it is *primarily* our problem but do we underestimate the extent to which a student's critical awareness of his course *as education* is itself a significant part of his education? Should his evaluation be merely in the role of consumer in the sense that he criticises what others provide and is in no significant sense responsible for what he consumes, nor responsible for giving an account of the relationship between what is provided and his own developing interests, enthusiasms, fears, inhibitions, abilities, sensitivities, pre-

judices, style of working and generally the sort of person he wants to become?

We want to suggest that in both the areas of formative student evaluation and course evaluation there are significant aspects which appear to be seriously under-emphasised. As yet we cannot point to evidence which will even clearly define the problems in these areas, much less point to effective ways of coping with them, but there is enough to alert us to the need for broadening our perspective.

THE MEANING OF DIFFERENT PATTERNS OF RESPONSE TO FORMATIVE STUDENT EVALUATION

We suggested that 'problems' arose in interpreting poor performance when lack of effort or ability were not clearly responsible. Of course, sometimes lack of effort is problematical but where diagnostic formative testing is so obviously useful is where the student is given information about the need for a redistribution of effort to areas of weakness, or a total increase in effort which he was not aware of before and which he can act on. Where he knows he is not working enough, wants to but cannot, then we have a tendency to call in the counsellor to help him with his 'emotional' problem which probably has something to do with sex, drugs, his family or perhaps anxiety about his future career. Where the diagnostic testing reveals a weakness which is not related to lack of effort it can still be directly useful if there is a real lack of ability which may be compensated for by disproportionate effort or some form of remedial teaching or learning - or even a change of course.

With many teachers lack of ability or effort exhaust the possibilities but many are aware that somewhere between 'the psychiatric', 'the cognitive' and the common-sense understanding of interest or motivation (effort) there is an area which is vitally important both to academic achievement in the narrower sense and to the possibility of education in the broader sense. It concerns attitudes, styles of work, strategies of learning, defences, expectations, adaptations, self-perceptions and many other rather vague concepts which are so easily relegated to no-man's land. But as assessment has become an almost universal focus for student discontent and as education has become less a socialisation or initiation process into adult conceived roles and expectations and more a period of exploration, of knowledge, of interests, of competencies and identity, then the need for a more personal understanding of the meaning of differential formative evaluation

becomes more important (Cox, 1973). Psychiatrists on both sides of the Atlantic have pointed to the way students in late adolescence can be working through unresolved family problems of authority and independence in relation to the demands of their colleges (Blaine & McArthur, 1961; Ryle, 1969). The rejection of assessment as a genuine measure of academic competence encourages the association of testing with authority and control rather than differential diagnosis. If student evaluation is to be truly formative today it needs to be seen within this broader understanding of the social context of education and especially in relation to the psychology of adolescence.

If we attempt to look at the meaning of student evaluation from this point of view what can we learn? We shall first discuss two illustrative individual cases then look at different responses to particular types of assessment and finally mention some research from Edinburgh on 'cue-consciousness'.

One of the most interesting cases in a study of student responses to assessment at Essex University was a chemistry student. He was one of the highest scorers on a test of divergent thinking and well above the average even of the arts students who were as usual much higher than the scientists. Other tests showed wider intellectual interests than average and no lack of general intelligence yet his general academic assessments were well below average and interviews exploring his style of work and approach in assessment revealed not a lack of effort but an extremely constricted and defensive way of responding to any academic threat. Although he could say 'I've always rather enjoyed exams. I enjoyed the chemistry exam tremendously; I knew the work and if you know the work they're fun' and although he adopted a rigid technique to maximise his marks and avoided any issues which were not safe he felt 'ability as a chemist was not in question in examinations'. 'Much of what I put into it is parrot learning. I suppose that's why chemistry appeals to me - that's my ability, a very good long-term memory.' What he felt he needed was control - 'something regular which can keep you on the lines you're supposed to be following'. 'I never really need a reason; I'm quite content to do it because it's there whereas other students, particularly those who have been in industry, like to know why they're doing things'. Asked for important features of a good essay his reply, after some hesitation, was 'one that rightly gets good marks'.

As might be expected at times he was profoundly unhappy about his approach. One essay gave him an opportunity to branch out - to actually express his broader interests and divergent ability but he said

sadly 'What I actually did was to stick to the course. It was a limited essay. I could have gone further afield and it would have been more beneficial but I tend not to write this sort of essay because it takes you outside your subject and I have always kept things in appropriate cubby holes'.

We have quoted at some length from this student because it seemed he was evaluating his own responses to assessment and to this subject in a way which, if it were pursued sympathetically by a tutor, might enable him to emerge from this own cubby hole and use his ability and develop interests which are so firmly held in check. A tutor rather than a counsellor may be able both to require the student to take more risks (as the researcher did with the divergent test) and to set up an intermediate stage of contact where he could discuss half-formed ideas and help to build them up with encouragement and constructive criticism. Another approach which might be suggested from this personal exploration of formative evaluation is the need for more group work where the authority of the tutor is lessened and the threat of formal individual evaluation may be absent. It is unlikely that the tutor would have the time or ability to understand the difficult and ambivalent relationship the student had with his father, but it seems very likely that much could be done without even attempting it.

Another student, studying sociology, reacted in a highly ambivalent way to authority - predominantly with aggression and hostility and sometimes with cynical conformity. He feared that in order to gain high marks he would need to 'distort (his) personality beyond recognition'. The family background to this student was extremely interesting but not essential to understand provided the tutor realised the strong need he had for both security and independence. He needed the security of winning approval and a sense of competence through gaining high marks in the dominant form of assessment, the traditional examinations; he needed independence through rejecting this form of assessment and working within what he took to be a genuine form of intellectual self-expression. It seemed that one was not possible without the other. He adopted a cynical but effective method of writing what he described as 'standard product' essays in examinations and this left him free to be himself in working on his project. He did not find that course essays (one every three weeks or so) created the right balance between his own expectations and the time available for working on them. Unlike the project the time was not long enough to muster sufficient confidence to deal with his higher expectations and the 'standard product' was not appropriate. The prior notice examinations,

however, where the questions were given out two weeks in advance did enable him to work out some of his own ideas while not raising his critical expectation too high.

If the tutor could manage to talk to this student without raising too much hostility he might be able to explore this very interesting but very individual area of balance between perceived opportunities. constraints and expectations, but perhaps more important he might be able to look at the interdependence between different expressions of ability. If we want risk-taking and creative work how far will it depend on creating a complementary area of secure - if pedestrian - competence? In this case did his area of security in the traditional exams involve too high a degree of cynicism and how serious were his feelings of distortion of personality? These issues seem to be important and yet they are not easy to explore without relaxed discussion of the meaning of different assessments with individual students.

Although the more significant problems concerning student formative evaluation may involve an understanding of its meaning in terms of the general pattern of the students response to education, there are other problems concerning the interpretation of grades. A particular form of assessment may need to be looked at very carefully if we are to be clear about how far it provided the opportunity for the expression of a particular type of ability. In the previous case we said the student found the examinations where the questions were given out two weeks in advance gave him the opportunity to work out some of his ideas so that he felt able to put more of his own thinking into his answers than with unseen questions. In terms of constraints the higher time and memory constraints of the unseen questions led him to restrict his performance to a 'standard product'.

With other students, however, exactly the opposite occurred since they were more affected by the constraints imposed by higher expectations on the prior-notice questions. They were afraid of the high standards so that they were unable to express their own ideas and looked for safety in reading more and more books. While *they* became less confident as the exam drew near, the others felt more confident that they would be able to present a reasonable case to support their ideas. Some students reported feeling more able to express their own ideas with the unseen questions where they knew very little could be expected of them. Of course this does not mean that we can make no inferences about the form of assessment and the type of ability measured. Although it would be wrong to assume that where prior-notice questions produced no evidence of original or critical thinking the student then lacked such ability, it might be reasonable to assume a certain

appropriateness but at the same time explore factors which prevent some students from using the opportunity. While this type of personalised approach does lead us to provide a wide range of opportunities for the expression of different abilities and a willingness to accept them wherever we find them, it focuses, too, on the different problems students face in coping with different types of constraint. The more bureaucratic type of assessment where the student's profile of abilities is formally linked to specific measuring situations may tend to obscure the significance of the less tangible factors in student performance.

There has been a tendency here to view 'high expectations' in a very general way, but of course it has many aspects associated with both the type of expectation and the source. 'Inner' and 'other' directedness can be a useful if crude distinction provided we realise that as with the sociology student it can operate in different ways in different areas. The type of expectation can of course vary both within the cognitive ability area and the broader areas of education, of values and life styles, and the mapping out of those expectations can be an invaluable guide to understanding the meaning of formative assessments. Miller and Parlett (1973) at Edinburgh University have shown that the way students approach the expectations of staff - in science at least - can affect their performances. They found they could divide students into three groups: the 'cue-seekers' who actively engaged in finding out about what might influence their chances in the assessments; the 'cue-conscious' who were perceptive and receptive to the cues sent out by staff but did not actively seek them; and the 'cue-deaf' - 'for them it seemed working as hard as you could was the ingredient for success'. They did not speak of picking up hints and did not think the impression they made on staff would affect their marks. The 'cue-seekers' performed better than the 'cue-conscious' who in turn did very much better than the 'cue-deaf'. The 'professionalisation' of examination-taking seemed to pay off. If evaluation is to be educationally formative in the best sense it needs to be seen within the broader context which takes such factors into account.

These types of study are of course not new. The Edinburgh work derived from Benson Snyder's (1971) studies of the hidden curriculum and earlier work by Mechanic (1962) and Becker, Geer and Hughes (1968) discussed assessment within the social context of student life. The resources and opportunities for this type of work are not open to tutors but the tutors have the great advantage of being able to change their formative evaluation and discuss the results with their students. Researchers, moreover, are generally not able to judge how far students'

perceptions are actually reflected in their work except in very crude ways.

THE USE OF COURSE EVALUATION AS A FORMATIVE INFLUENCE ON STUDENTS

Earlier we asked the question whether course evaluation should be so predominantly the teacher's concern, with the student's role confined to providing fairly low-level information and opinions. In Britain, we ask students to be deeply critical and concerned about many things but not about their most central intellectual and often emotional concern - their own education. This is significant now for many reasons. Perhaps most important is the question of structure. Rigidly structured courses which impose strict limits on choice and the opportunity to devise individual learning experiences are becoming suspect. Styles of teaching which are didactic or authoritarian are questioned. Knowledge itself is seen as more open to revision, more interdisciplinary and less hierarchical. These changes are not independent of broader issues of control, authority and participation in society nor of adolescents' concern for identity and the achievement of independence. In the past we could compartmentalise the authority relationships and structuring of the course so that contradictions between what might crudely be described as means and ends were not apparent. This is less so today but the ambivalence students feel towards assessment is apparent too in relation to their learning experiences where they may reject control yet still fear freedom. In an evaluation study, carried out by our unit, of a course for teachers in Medical General Practice it became clear that the doctors' involvement in evaluating the course enabled them to relate the structure of the learning experiences to both their need for independence as experienced professional practitioners, and their need for structure which was derived from their previous formal educational experience and unfamiliarity with new and often alien ways of thinking. It enabled them, too, to experience and work through aspects of group interaction which was important in helping them to use what they had learned.

This course was set up in London in 1970 by the Royal College of General Practitioners and the British Postgraduate Medical Federation. We were invited to attend the 24 weekly meetings with a view to devising new methods of evaluating the course. In the autumn of 1971 funds were granted by the Department of Health and Social Security to provide a medically qualified part-time research fellow to assist in the evaluation during the new courses in 1972 and 1973.

Originally five kinds of activity were planned:

1. (a) Elaboration and clarification of a scheme of objectives and activities for trainees; and (b) development of a scheme of objectives for the course itself.

2. Use of audio tapes of interviews between trainers and their trainees in order to identify any changes in teaching behaviour as a consequence of the course.

3. Recording on video tape of (a) teaching behaviour in role-playing interviews with 'patients' and subsequent discussions between trainer and trainee; and (b) presentations of case studies by course members.

4. Course members (a) were invited to fill in short questionnaires indicating their opinions of each session of the course; (b) discussed the course as a group in formal and spontaneous sessions; and (c) at the end were interviewed about the course individually.

5. Completion, before and after the course, of an attitude questionnaire.

The question of structure became a dominant theme in all the course evaluation discussions, in the sessional evaluation questionnaire and the end-of-course interviews, and there were interesting contrasts between the three courses. The first course - before we became fully involved - began by being highly structured (lecture followed by short discussion repeated four times in each afternoon). After five sessions of this there was a near revolution which threatened the existence of the course, but after long discussions about what the course was trying to do it developed a much looser structure which worked well with a high level of commitment. The organisers thought they had learned their lesson from this and began the next course with a loose structure to explore what the course members wanted. After a few weeks extreme hostility was provoked in some members and sympathy for this in many others. Eventually a somewhat more structured programme developed and the course worked well with a high level of commitment - except for the few who dropped out. Again, the organisers thought they had learned their lesson and the third course began with carefully balanced elements of structure and openness set in a residential weekend in a hotel in Brighton, so that relationships could be established and problems worked through both formally and informally. But has this therefore been the most successful course? Certainly we have discussed questions of authority, structure and independence and they seem to say the 'right' things. As one member said in a discussion with the previous year's members, 'If you do not want purely didactic teaching then you need a woolly stage at the beginning'.

The involvement of the members in the evaluation of the course - not as consumers but as participants who have a responsibility for making it into the sort of course they want, and the sort of course which makes sense

course at Sussex undergraduates spend their last two years working entirely on research projects.

Problems in evaluation arise because of the different amounts of assistance students receive from tutors or members of a research team, from the difficulty of a topic, and from the competence of tutors and examiners to guide and to assess the work. In the end, similar grades are assigned to essentially dissimilar pieces of work having different emphases on such aspects as discussion of relevant literature, complexity of the problems solved, elegance of experimental design, literary skill in writing the thesis, and in originality. It is becoming more important, therefore, both to explore individual responses and involve the students more. From a study of group projects where tutors set up the problems then acted as consultants it seemed students welcomed greater interest in the processes of education. Most felt that it would be helpful to add a section to the project report on how useful it was as a learning exercise. Diaries in which they could keep notes on activities and their own personal responses to them were thought to be worth keeping on future courses but even without these the project evaluation discussions had led them to be more critical of their own education. This may have been more effective had they been more involved in actually awarding grades, this is not uncommon with architectural design projects. The value of such self- or peer-assessment is not only that the students become more aware and more able to use the criteria of critical judgment but the assessment is more firmly anchored in the characteristics of the work rather than the authority of the teacher.

This approach to formative evaluation which attempts to actively involve students and explores the meaning of processes, activities and performances in terms of the wide variety of students' perceptions and responses is in line with recent general thinking about evaluation. Stake (1967) and Pace (1972) in America, and Parlett (1972) now in Britain, have advocated a more open approach which can take more account of the unexpected and the less clearly defined, and concern for the college as Pace sees it, 'an environment for exploration - of self, of knowledge and skills, of ideas and values, of society, conscience, community and commitment'. The analogies are no longer dominated by the mechanical 'not just' - 'What does your machine produce?' but also 'How does your garden grow?'

in terms of their future teaching roles - has given them an experience of constructive criticism and educational change which we hope will now be a natural part of their approach to their own teaching. Our fear is that while it is based upon experience, that experience has been rather too stage-managed. The present members came away from a joint meeting feeling that the previous course had had something which they lacked and which was significant. As one put it 'The woolly period was not long enough for us, we haven't had time to work it through'. But the course was almost over. What they lacked was a genuine feeling that the future of the course really was dependent upon their working through conflicts to produce a solution which was their own. The difference in tone was clear when one of the most intellectual members of the previous course and one of the more hostile early critics summed up their initial expectations of the course, and, implicitly, the distance they had travelled, 'We wanted Daddy to give us the goodies'.

This particular course perhaps emphasised the social psychology of the learning context and short questionnaires on group interaction helped them to learn from it, but it is part of all courses. We believe that formative evaluation of both the individual responses to the learning experience - especially in relation to demands for structure, control and openness - and the significance of group relationships and attitudes can be a productive link to reality. Students often complain of artificiality in their lives - so much of what they do and think about is at one remove from reality, part of an artificially contrived learning environment - but their education itself is real and the responses of teachers and students change it, they do not merely simulate change. Yet so often student thinking about their courses is perfunctory - ten-second-off-the-cuff responses about something which is over and done with as far as they are concerned. As long as course design is seen so exclusively as the job of the teacher the degree of involvement is likely to remain limited and, indeed, inhibited. Slowly, however, aspects of courses are becoming areas of student initiative where staff act as consultants. Here it is not a staff course evaluated by a consumer student but a student course evaluated by staff or rather by staff and students together. Emphasis, however, still tends to be on student products rather than educational processes. But conventional evaluation of products presents considerable difficulties where teaching methods consist of projects or experimental work to solve genuine problems. This has always been the case of students presenting theses to gain higher degrees, but is now more common in undergraduate courses. In London, undergraduates in engineering courses spend some 50 per cent of their time working on projects which include genuine problems from industry; and in the 'Eaborn' chemistry degree

REFERENCES

Becker H S, Geer B & Hughes E (1968) *Making the grade.* New York: Wiley.

Blaine G & McArthur C (1961) *The emotional problems of the student.* New York: Appleton Century-Crofts.

Cox R J (1973) Traditional examinations in a changing society. *Universities Quarterly 27(2)*,200-16.

Mechanic D (1962) *Students under stress.* New York: Free Press.

Miller C & Parlett M (1973) *Up to the mark: A research report on assessment.* Edinburgh: Centre for Research in Educ. Sciences. Edinburgh University Occasional Paper.

Pace C R (1972) *Thoughts on evaluation in higher education.* Iowa City: Amer. Coll. Testing Prog.

Parlett M (1972) Evaluating innovations in teaching. In *Contemporary problems in higher education*, Butcher H & Rudd E (eds.), London: McGraw-Hill.

Ryle A (1969) *Student casualties.* London: Allen Lane.

Snyder B (1971) *The hidden curriculum.* New York: Knopf.

Stake R E (1967) The countenance of educational evaluation. *Teach. Coll. Rec, 68,* 523-540.